



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 01:46 am GMT

PDB ID : 4RYP  
Title : Crystal Structure of T. Brucei Farnesyl Diphosphate Synthase  
Authors : Cao, R.; Liu, Y.-L.; Oldfield, E.  
Deposited on : 2014-12-16  
Resolution : 2.21 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

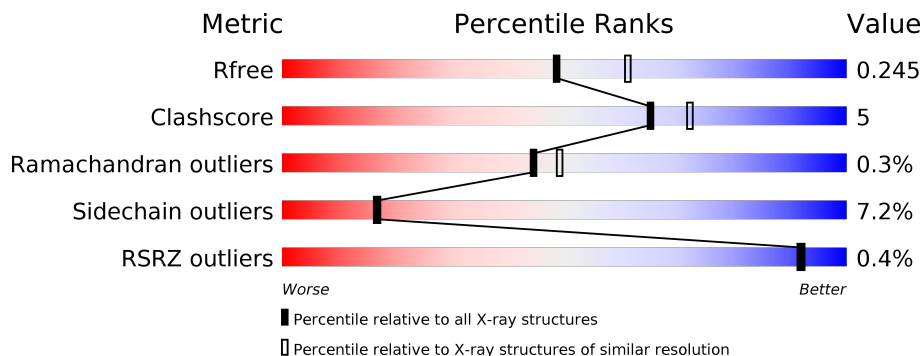
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.21 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	4744 (2.24-2.20)
Clashscore	112137	5509 (2.24-2.20)
Ramachandran outliers	110173	5427 (2.24-2.20)
Sidechain outliers	110143	5428 (2.24-2.20)
RSRZ outliers	101464	4776 (2.24-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	390	<div> <div> <div></div> <div>71%</div> <div>17%</div> <div>•</div> <div>10%</div> </div> </div>
1	B	390	<div> <div> <div></div> <div>75%</div> <div>13%</div> <div>•</div> <div>11%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5771 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Farnesyl pyrophosphate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	351	Total	C	N	O	S	0	0	0
			2811	1790	462	531	28			
1	B	347	Total	C	N	O	S	0	0	0
			2784	1775	457	524	28			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MET	-	EXPRESSION TAG	UNP Q86C09
A	-21	GLY	-	EXPRESSION TAG	UNP Q86C09
A	-20	SER	-	EXPRESSION TAG	UNP Q86C09
A	-19	SER	-	EXPRESSION TAG	UNP Q86C09
A	-18	HIS	-	EXPRESSION TAG	UNP Q86C09
A	-17	HIS	-	EXPRESSION TAG	UNP Q86C09
A	-16	HIS	-	EXPRESSION TAG	UNP Q86C09
A	-15	HIS	-	EXPRESSION TAG	UNP Q86C09
A	-14	HIS	-	EXPRESSION TAG	UNP Q86C09
A	-13	HIS	-	EXPRESSION TAG	UNP Q86C09
A	-12	SER	-	EXPRESSION TAG	UNP Q86C09
A	-11	SER	-	EXPRESSION TAG	UNP Q86C09
A	-10	GLY	-	EXPRESSION TAG	UNP Q86C09
A	-9	LEU	-	EXPRESSION TAG	UNP Q86C09
A	-8	VAL	-	EXPRESSION TAG	UNP Q86C09
A	-7	PRO	-	EXPRESSION TAG	UNP Q86C09
A	-6	ARG	-	EXPRESSION TAG	UNP Q86C09
A	-5	GLY	-	EXPRESSION TAG	UNP Q86C09
A	-4	SER	-	EXPRESSION TAG	UNP Q86C09
A	-3	HIS	-	EXPRESSION TAG	UNP Q86C09
A	-2	MET	-	EXPRESSION TAG	UNP Q86C09
A	-1	ALA	-	EXPRESSION TAG	UNP Q86C09
A	0	SER	-	EXPRESSION TAG	UNP Q86C09
B	-22	MET	-	EXPRESSION TAG	UNP Q86C09
B	-21	GLY	-	EXPRESSION TAG	UNP Q86C09

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-20	SER	-	EXPRESSION TAG	UNP Q86C09
B	-19	SER	-	EXPRESSION TAG	UNP Q86C09
B	-18	HIS	-	EXPRESSION TAG	UNP Q86C09
B	-17	HIS	-	EXPRESSION TAG	UNP Q86C09
B	-16	HIS	-	EXPRESSION TAG	UNP Q86C09
B	-15	HIS	-	EXPRESSION TAG	UNP Q86C09
B	-14	HIS	-	EXPRESSION TAG	UNP Q86C09
B	-13	HIS	-	EXPRESSION TAG	UNP Q86C09
B	-12	SER	-	EXPRESSION TAG	UNP Q86C09
B	-11	SER	-	EXPRESSION TAG	UNP Q86C09
B	-10	GLY	-	EXPRESSION TAG	UNP Q86C09
B	-9	LEU	-	EXPRESSION TAG	UNP Q86C09
B	-8	VAL	-	EXPRESSION TAG	UNP Q86C09
B	-7	PRO	-	EXPRESSION TAG	UNP Q86C09
B	-6	ARG	-	EXPRESSION TAG	UNP Q86C09
B	-5	GLY	-	EXPRESSION TAG	UNP Q86C09
B	-4	SER	-	EXPRESSION TAG	UNP Q86C09
B	-3	HIS	-	EXPRESSION TAG	UNP Q86C09
B	-2	MET	-	EXPRESSION TAG	UNP Q86C09
B	-1	ALA	-	EXPRESSION TAG	UNP Q86C09
B	0	SER	-	EXPRESSION TAG	UNP Q86C09

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	94	Total O 94 94	0	0
2	B	82	Total O 82 82	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.33Å 119.44Å 62.04Å 90.00° 117.18° 90.00°	Depositor
Resolution (Å)	29.87 – 2.21 29.87 – 2.21	Depositor EDS
% Data completeness (in resolution range)	96.2 (29.87-2.21) 94.7 (29.87-2.21)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.96 (at 2.22Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.180 , 0.241 0.187 , 0.245	Depositor DCC
$R_{free}$ test set	2104 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	42.0	Xtriage
Anisotropy	0.100	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 28.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.008 for k+l,h+l,-l 0.008 for -k+l,-h-l,-l 0.467 for -h-2*l,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5771	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.52% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.79	0/2867	0.89	8/3875 (0.2%)
1	B	0.80	0/2839	0.86	3/3835 (0.1%)
All	All	0.80	0/5706	0.88	11/7710 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	7
1	B	0	4
All	All	0	11

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	31	ARG	NE-CZ-NH2	-9.75	115.42	120.30
1	A	31	ARG	NE-CZ-NH1	8.48	124.54	120.30
1	A	50	ARG	NE-CZ-NH1	8.10	124.35	120.30
1	A	50	ARG	NE-CZ-NH2	-7.85	116.37	120.30
1	B	207	ARG	NE-CZ-NH2	-6.74	116.93	120.30
1	A	196	ASP	N-CA-C	5.80	126.67	111.00
1	A	103	ASP	CB-CG-OD1	5.73	123.45	118.30
1	A	196	ASP	CB-CA-C	-5.53	99.35	110.40
1	B	207	ARG	NE-CZ-NH1	5.52	123.06	120.30
1	B	50	ARG	NE-CZ-NH2	-5.29	117.66	120.30
1	A	40	ASP	CB-CG-OD1	5.25	123.03	118.30

There are no chirality outliers.

All (11) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	113	ARG	Peptide
1	A	114	GLY	Peptide
1	A	123	ASP	Peptide
1	A	193	THR	Peptide
1	A	195	THR	Peptide
1	A	269	LYS	Peptide
1	A	75	GLY	Peptide
1	B	196	ASP	Peptide
1	B	261	PHE	Peptide
1	B	269	LYS	Peptide
1	B	364	LYS	Peptide

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2811	0	2763	29	0
1	B	2784	0	2737	23	0
2	A	94	0	0	1	0
2	B	82	0	0	1	0
All	All	5771	0	5500	52	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (52) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:195:THR:HG22	1:A:197:PHE:H	1.47	0.79
1:B:106:MET:HE2	1:B:172:GLN:HE21	1.52	0.74
1:B:286:PHE:HB2	1:B:322:LEU:HD11	1.73	0.69
1:A:365:ARG:O	1:A:366:GLN:HG3	1.93	0.68
1:B:106:MET:CE	1:B:172:GLN:HE21	2.07	0.66
1:B:83:HIS:HD2	2:B:415:HOH:O	1.78	0.66
1:A:196:ASP:HB3	1:A:198:ALA:H	1.59	0.66
1:A:286:PHE:HB2	1:A:322:LEU:HD11	1.82	0.62
1:A:83:HIS:HD2	2:A:411:HOH:O	1.82	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:365:ARG:O	1:A:366:GLN:CG	2.49	0.61
1:B:342:GLU:OE2	1:B:345:ARG:NH1	2.36	0.58
1:A:57:VAL:O	1:A:61:LEU:HD23	2.05	0.56
1:A:301:ASN:HD22	1:A:309:LYS:HA	1.70	0.56
1:A:50:ARG:CZ	1:A:96:GLN:HE22	2.19	0.56
1:B:50:ARG:CZ	1:B:96:GLN:HE22	2.19	0.56
1:B:301:ASN:HD22	1:B:309:LYS:HA	1.70	0.55
1:A:98:HIS:HD2	1:A:133:GLY:O	1.90	0.54
1:B:141:HIS:HB3	1:B:145:MET:CE	2.38	0.54
1:A:180:PHE:HA	1:A:196:ASP:HB2	1.92	0.52
1:B:44:LEU:O	1:B:113:ARG:NH2	2.43	0.51
1:A:111:THR:HG22	1:A:116:PRO:HA	1.91	0.51
1:A:12:GLU:OE2	1:A:83:HIS:HE1	1.92	0.51
1:A:260:CYS:O	1:A:314:ARG:NH2	2.43	0.51
1:B:98:HIS:HD2	1:B:133:GLY:O	1.94	0.50
1:A:357:THR:O	1:A:361:LYS:HG2	2.15	0.46
1:A:323:GLN:O	1:A:327:VAL:HG23	2.16	0.46
1:A:258:MET:O	1:A:262:THR:HG22	2.16	0.46
1:B:196:ASP:CB	1:B:198:ALA:H	2.28	0.45
1:B:277:ALA:HB2	1:B:303:GLY:HA3	1.97	0.45
1:A:277:ALA:HA	1:A:302:TYR:CE2	2.53	0.44
1:A:342:GLU:OE2	1:A:345:ARG:NH2	2.48	0.44
1:A:366:GLN:HB3	1:A:366:GLN:HE21	1.69	0.43
1:A:237:MET:O	1:A:241:GLU:HG3	2.18	0.43
1:B:277:ALA:HB2	1:B:303:GLY:CA	2.49	0.43
1:A:100:LEU:HD22	1:A:112:ARG:NH2	2.34	0.43
1:B:12:GLU:OE2	1:B:83:HIS:HE1	2.01	0.43
1:B:232:LEU:N	1:B:233:PRO:CD	2.81	0.43
1:A:179:MET:O	1:A:195:THR:HG23	2.19	0.42
1:B:10:TYR:HB2	1:B:90:TRP:CZ2	2.53	0.42
1:B:141:HIS:HB3	1:B:145:MET:HE3	2.01	0.42
1:B:141:HIS:HB3	1:B:145:MET:HE1	2.00	0.42
1:A:179:MET:O	1:A:195:THR:HA	2.20	0.42
1:B:237:MET:O	1:B:241:GLU:HG3	2.20	0.42
1:B:100:LEU:HG	1:B:112:ARG:CZ	2.49	0.41
1:B:277:ALA:HA	1:B:302:TYR:CE2	2.56	0.41
1:A:10:TYR:HB2	1:A:90:TRP:CZ2	2.55	0.41
1:B:189:VAL:HG12	1:B:190:SER:O	2.21	0.41
1:B:18:LEU:HA	1:B:18:LEU:HD23	1.90	0.41
1:A:125:THR:O	1:A:128:CYS:N	2.53	0.41
1:A:232:LEU:N	1:A:233:PRO:CD	2.84	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:108:ASN:O	1:A:108:ASN:CG	2.60	0.40
1:A:218:TYR:CD2	1:A:248:GLY:HA2	2.57	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	345/390 (88%)	331 (96%)	12 (4%)	2 (1%)	28	28
1	B	339/390 (87%)	327 (96%)	12 (4%)	0	100	100
All	All	684/780 (88%)	658 (96%)	24 (4%)	2 (0%)	44	49

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	LEU
1	A	126	VAL

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	306/340 (90%)	282 (92%)	24 (8%)	15	14
1	B	303/340 (89%)	283 (93%)	20 (7%)	19	20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	609/680 (90%)	565 (93%)	44 (7%)	17	17

All (44) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	LYS
1	A	26	ASP
1	A	80	ARG
1	A	82	LEU
1	A	112	ARG
1	A	120	ARG
1	A	125	THR
1	A	182	SER
1	A	190	SER
1	A	195	THR
1	A	203	SER
1	A	216	TYR
1	A	243	LEU
1	A	269	LYS
1	A	282	LEU
1	A	307	SER
1	A	308	GLU
1	A	318	GLU
1	A	338	LYS
1	A	340	LEU
1	A	343	LYS
1	A	346	LEU
1	A	361	LYS
1	A	366	GLN
1	B	1	MET
1	B	24	LYS
1	B	61	LEU
1	B	74	ASP
1	B	78	ARG
1	B	100	LEU
1	B	125	THR
1	B	135	LEU
1	B	182	SER
1	B	191	GLN
1	B	195	THR
1	B	203	SER
1	B	207	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	216	TYR
1	B	243	LEU
1	B	262	THR
1	B	269	LYS
1	B	270	VAL
1	B	340	LEU
1	B	346	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	ASN
1	A	83	HIS
1	A	96	GLN
1	A	98	HIS
1	A	183	ASN
1	A	191	GLN
1	A	204	ASN
1	A	254	GLN
1	A	301	ASN
1	A	366	GLN
1	B	30	ASN
1	B	83	HIS
1	B	96	GLN
1	B	98	HIS
1	B	183	ASN
1	B	204	ASN
1	B	254	GLN
1	B	301	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	351/390 (90%)	-0.43	2 (0%) 89 88	30, 47, 85, 117	0
1	B	347/390 (88%)	-0.43	1 (0%) 93 93	31, 47, 83, 121	0
All	All	698/780 (89%)	-0.43	3 (0%) 92 92	30, 47, 85, 121	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	270	VAL	2.3
1	A	270	VAL	2.2
1	A	114	GLY	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.